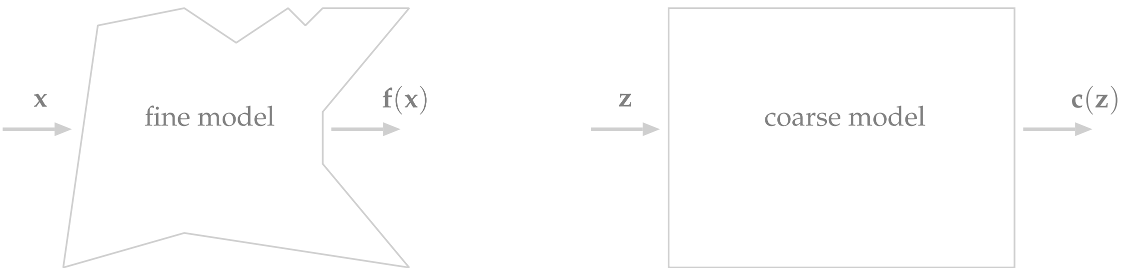
# Space Mapping Optimization and Model Reduction for Biogeochemical Models

A3: Algorithmic Optimal Control - C02 Uptake of the Ocean

(Prof. Dr. Thomas Slawig)




Malte Prieß ◦ mpr@informatik.uni-kiel.de

Institut für Informatik

IFM-GEOMAR
Leibniz-Institut für Meereswissenschaften an der Universität Kiel

ifw

future ocean
KIEL MARINE SCIENCES

# Cluster of Excellence “The Future Ocean”

CAU
Christian-Albrechts-Universität zu Kiel

muthesius kunsthochschule

# General Research Aims

- The ocean ↷ key role in the climate system
- Already suffering from global warming
- Cover more than two thirds of our planet but yet little explored

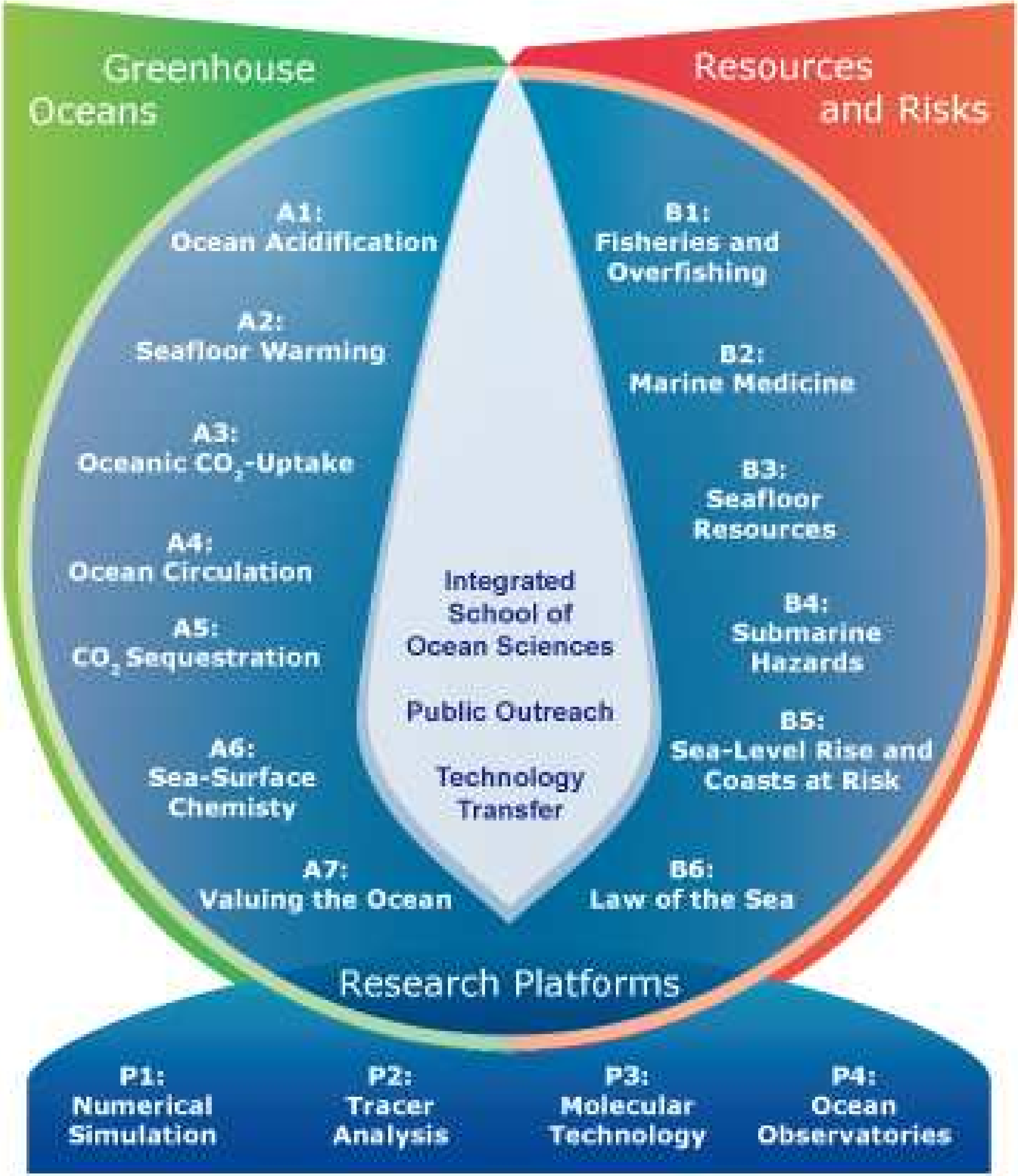
The Aim:

Looking at past, present, future ocean changes

Investigating marine ressources

Developing techniques for their sustainable use

↷ Increase our understanding of ocean change and its potential and risks


Greenhouse Oceans
Resources and Risks
A1: Ocean Acidification
A2: Seafloor Warming
A3: Oceanic $CO_2$-Uptake
A4: Ocean Circulation
A5: $CO_2$ Sequestration
A6: Sea-Surface Chemisty
A7: Valuing the Ocean
B1: Fisheries and Overfishing
B2: Marine Medicine
B3: Seafloor Resources
B4: Submarine Hazards
B5: Sea-Level Rise and Coasts at Risk
B6: Law of the Sea
Integrated School of Ocean Sciences
Public Outreach
Technology Transfer
Research Platforms
P1: Numerical Simulation
P2: Tracer Analysis
P3: Molecular Technology
P4: Ocean Observatories

# Algorithmic Optimal Control - C02 Uptake of the Ocean



- Natural causes + anthropogenic $CO_2$ emissions $\curvearrowright$ global warming
- $CO_2$ concentration has doubled since 1900
- To-date we assume $4-8°$C in the buisiness as usual case
- Agreement on the “2-degree-aim” until the year 2100
- This relates to a $CO_2$ emission reduction about 80% until 2050 (w.r.t. 1990)
- Concentrating only on a sustainable energypolitics will not comply with this aim
- Moreover we need to strongly think of carbon management/ sequestration approaches

- The ocean $\curvearrowright$ biggest $CO_2$ sink

  More than half of anthropogenic $CO_2$ stored for long time

  $\curvearrowright$ Crucial impact on climate

- Natural Sequestration based upon global $CO_2$ cycle

- “Physical + Biological $CO_2$ pump” are the operators

  $\curvearrowright$ $CO_2$ can remain in the deep sea for years

- Ocean Circulation + Biogeochemical Models indispensible

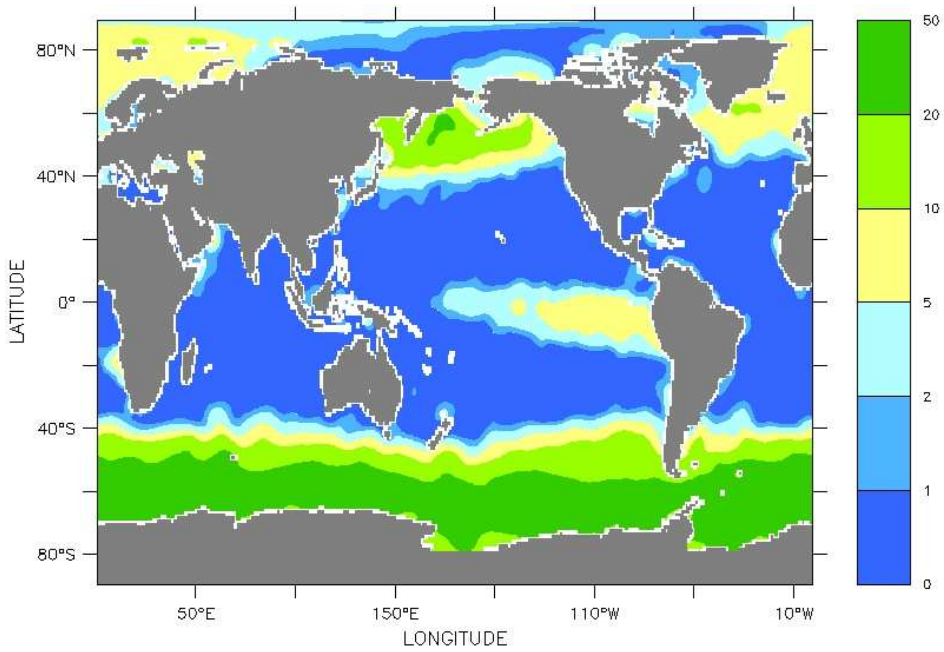
Reserach Aims:

Reduce large uncertainties in existing biological models

$\curvearrowright$ improve determination of current/ future $CO_2$ sequestration potential

# The Biogeochemical Models

# Motivation




Present-day sea-surface nitrate concentrations (Conkright et al., 1994)

---

- Represent ecological processes contributing to global C02 cycle
- Various models differing in complexity (# of state variables)
- Available data places significant limitations on complexity
- Nitrogen-based ecosystem model ↷ standard model

↷ 0-D transport-only to fully 3-D offline/ online coupled physical-biological simulations

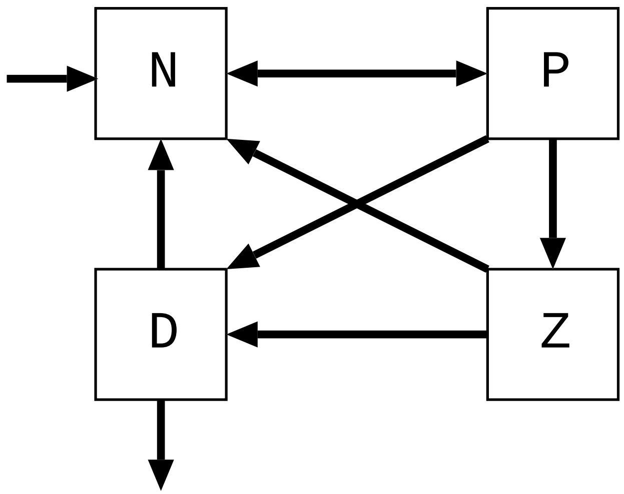
# The Model Equations

- Linear transport/ advection – diffusion eqs. with nonlinear forcing $q_i$

$$\frac{\partial y_i}{\partial t} = \underbrace{-\mathbf{v}\,(\nabla\, y_i)}_{\text{advection}} + \underbrace{\nabla(\kappa\,\nabla\, y_i)}_{\text{diffusion/ mixing}} + \underbrace{q_i(\mathbf{y},\mathbf{u},t)}_{\text{biological processes}}$$

- “Real world” simulation: coupling to ocean circulation models via the velocity field $\mathbf{v}$ necessary (offline via TMM, online)

$$\begin{array}{lll} \mathbf{y} & \in & \{N,P,Z,D\} \in \mathbb{R}^{\sim 10^7} \\ \mathbf{u} & \in & \mathbb{R}^{12} \end{array}, \quad \left\{ \begin{array}{lll} N = N(P,Z,D) & : & \text{dissolved inorganic nitrogen} \\ P = P(N,Z) & : & \text{phytoplankton} \\ Z = Z(P) & : & \text{zooplankton} \\ D = D(P,Z) & : & \text{detritus} \end{array} \right.$$



Institut für Informatik

Christian-Albrechts-Universität zu Kiel

future ocean
KIEL MARINE SCIENCES

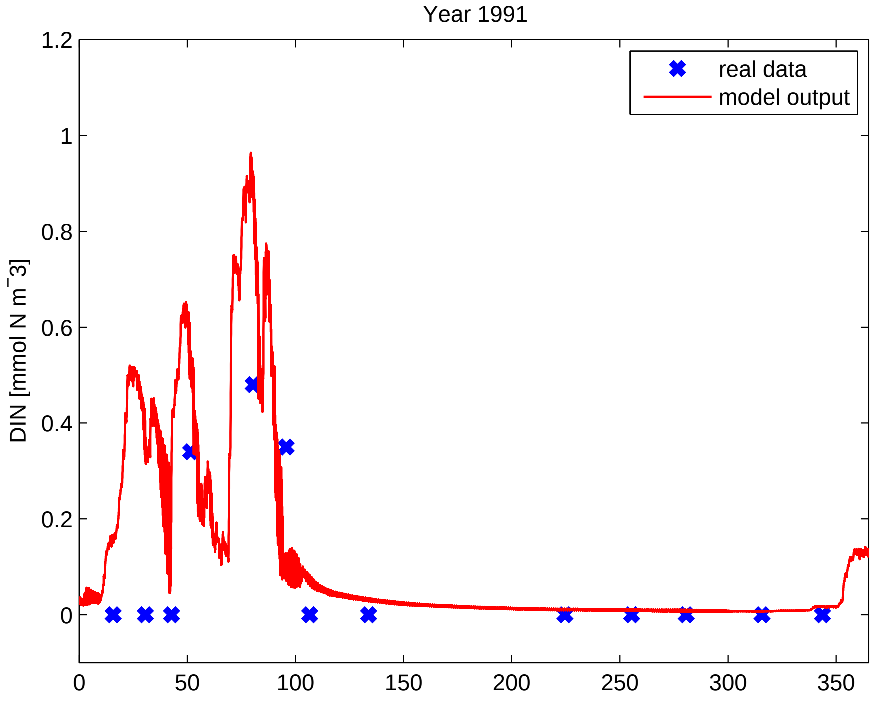
# The Optimization (OPT) Problem

- Minimize distance between the model output $\mathbf{y}(\mathbf{u}_f)$ and the desired state $\mathbf{y}_d$ (obs. data)

$$\text{argmin}_{(\mathbf{y},\mathbf{u}_f)}\ \mathcal{J}(\mathbf{y},\mathbf{u}_f) \quad , \quad \mathcal{J}(\mathbf{y},\mathbf{u}_f) = \left[ \frac{1}{2} \cdot \left\| \mathbf{y} - \mathbf{y}_d \right\|^2 + \frac{\alpha}{2} \cdot \left\| \mathbf{u}_f - \bar{\mathbf{u}}_f \right\|^2 \right]$$

$$\text{s.t.} \qquad e(\mathbf{y},\mathbf{u}_f) = 0 \quad ; \quad \mathbf{u}_l \leq \mathbf{u} \leq \mathbf{u}_u$$

- Control variables are the unknown physical/ biological parameters $\mathbf{u}$ in the nonlinear coupling terms ( $\mathbf{u}$ stationary in time and space! )

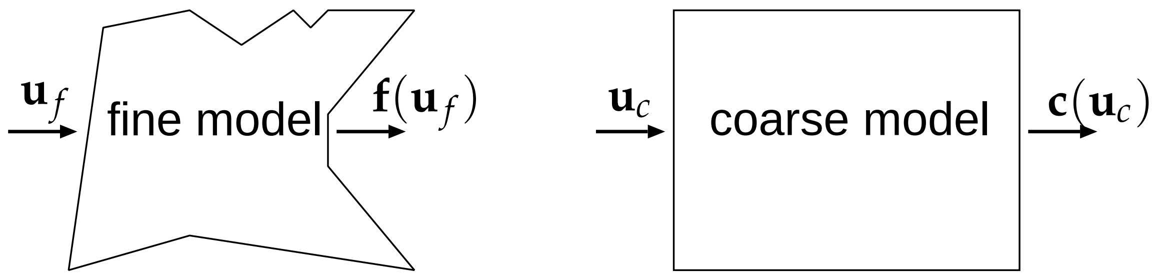


# Space Mapping (SM) Optimization

# Aims and First Definitions



- SM approach quit succesfully applied for engineering models so far
- Seek at optimum of complex “fine” model
- SM drives the OPT of the fine model to a fast “coarse” model
  - ↷ Coarse model shares the same physics as the fine counterpart
  - ↷ avoiding compuationally expensive fine model gradients and evaluations
- Key element is the mapping function (essential subproblem )
- Crucially depends on model similarity/ discrepancy
- Focus lies on the development of appropiate coarse models




## fine model

| | | |
|---|---|---|
| $\mathbf{u}_f^* = \text{argmin}_{\mathbf{u}_f} H_f(\mathbf{u}_f) := \frac{1}{2}\left\|\mathbf{f}(\mathbf{t},\mathbf{u}_f) - \mathbf{y}\right\|^2$ | : | fine model optimum |
| $\mathbf{u}_f \in \Omega_f \subset \mathbb{R}^{n_f}$ | : | control parameters |
| $\mathbf{y}_d \in \mathbb{R}^m$ | : | desired state |

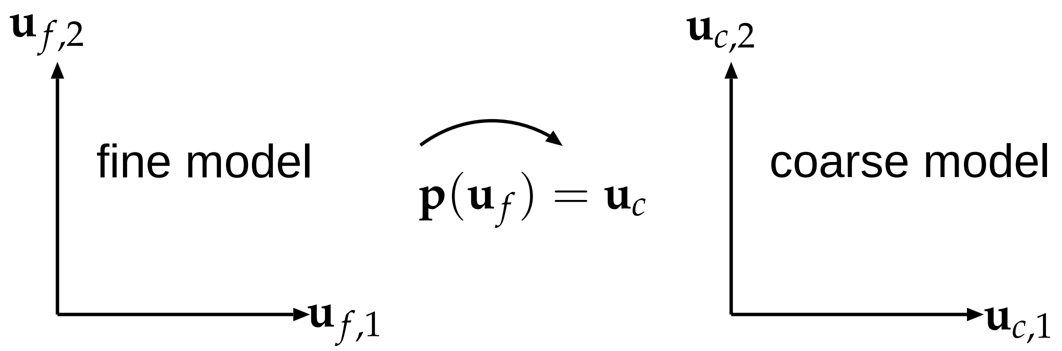
$\curvearrowright$ accurate but expensive, derivatives expensive/ not available

## coarse model

| | | |
|---|---|---|
| $\mathbf{u}_c^* = \text{argmin}_{\mathbf{u}_c} H_c(\mathbf{u}_c) := \frac{1}{2}\left\|\mathbf{c}(\mathbf{t},\mathbf{u}_c) - \mathbf{y}\right\|^2$ | : | coarse model optimum |
| $\mathbf{u}_c \in \Omega_c \subset \mathbb{R}^{n_c}$ | : | control parameters |

$\curvearrowright$ less accurate but fast, derivatives cheap

# The SM Function




SM establishes mapping $\mathbf{p} : \Omega_f \to \Omega_c$ s.t.

$$\mathbf{f}(\mathbf{u}_f) \simeq \mathbf{c}\big[\,\mathbf{p}(\mathbf{u}_f)\,\big]$$

$\mathbf{p}$ is defined as

$$\mathbf{u}_c = \mathbf{p}(\mathbf{u}_f) = \text{argmin}_{\hat{\mathbf{u}}_c \in \Omega_c} \overbrace{r(\hat{\mathbf{u}}_c, \mathbf{u}_f)}^{\text{misalignment function}} \quad , \quad r(\mathbf{u}_c, \mathbf{u}_f) = \frac{1}{2} \left\| \mathbf{c}(\mathbf{u}_c) - \mathbf{f}(\mathbf{u}_f) \right\|^2$$

Now, replacing $\mathbf{f}$ by its surrogate $\mathbf{c} \circ \mathbf{p}$ , we obtain two SM approaches

$$\underbrace{\bar{\mathbf{u}}_f^{(d)} = \text{argmin}_{\mathbf{u}_f \in \Omega_f} \frac{1}{2} \left\| \mathbf{c}\big[\,\mathbf{p}(\mathbf{u}_f)\,\big] - \mathbf{y} \right\|^2}_{\text{dual SM approach}} \quad \overset{(*)}{\Longleftrightarrow} \quad \underbrace{\mathbf{p}(\bar{\mathbf{u}}_f^{(p)}) - \mathbf{u}_c^* = 0}_{\text{primal SM approach}}$$

$(*)$ Only satisfied under certain “ideal” conditions

# Ideal Conditions

Let $\bar{U}_f$ and $\bar{U}_c$ be the sets of all SM solutions and coarse model minimizers

**C1** $\bar{U}_c \subseteqq \mathbf{p}(\mathbb{R}^{n_c})$

**C2** $\bar{U}_c \subseteqq \mathbf{p}(\bar{U}_f)$ (perfect mapping )

**C3** $\mathbf{p}$ is injective

**C4** $\bar{U}_f$ and $\bar{U}_c$ are singletons

If conditions **C1** - **C4** hold we yield the ideal case :

$$\bar{\mathbf{u}}_f^{(p)} = \bar{\mathbf{u}}_f^{(d)} = \mathbf{u}_f^* \quad ; \quad \mathbf{p}(\mathbf{u}_f^*) = \mathbf{u}_c^*$$

# Example: The Aggressive SM (ASM) Algorithm



ASM just solves the primal SM problem

$$\mathbf{F}(\mathbf{u}_f) = \mathbf{p}(\mathbf{u}_f) - \mathbf{u}_c^* \stackrel{!}{=} 0 \qquad (1)$$

by a quasi-Newton iteration and a Broyden rank-one update

$$\mathbf{u}_f^{(k+1)} = \mathbf{u}_f^{(k)} + \mathbf{s}^{(k)} \quad , \quad B^{(k)}\ \mathbf{s}^{(k)} = -\mathbf{F}^{(k)}$$

$$B^{(k+1)} = B^{(k)} + \frac{\mathbf{F}^{(k+1)}\ {\mathbf{s}^{(k)}}^T}{\mathbf{s}^{(k)}\ {\mathbf{s}^{(k)}}^T} \quad , \quad \mathbf{F}^{(k)} := \mathbf{F}(\mathbf{u}_f^{(k)}) = \mathbf{p}(\mathbf{u}_f^{(k)}) - \mathbf{u}_c^*$$

Each step requires the evaluation of $\mathbf{p}$, hence one fine model evaluation

$$\mathbf{p}(\mathbf{u}_f^{(k)}) = \text{argmin}_{\mathbf{u}_c} \frac{1}{2} \left\| \mathbf{c}(\mathbf{u}_c) - \mathbf{f}(\mathbf{u}_f^{(k)}) \right\|^2$$

More conveniently (1) is often using the least-sqare formulation

$$\bar{\mathbf{u}}_f = \text{argmin}_{\mathbf{u}_f} \left\| \mathbf{F}(\mathbf{u}_f) \right\|^2 \quad ; \quad \mathbf{F}(\mathbf{u}_f^k + \mathbf{s}^k) \simeq \mathbf{F}(\mathbf{u}_f^k) + \mathbf{B}^{(k)}\ \mathbf{s}^k$$

# The Coarse Models so Far

# Coarse-Discretization Model

- Coarse-discretized (in time) version based upon the same model
- OPT of function $r(\mathbf{u}_c, \mathbf{u}_f)$ obtaining mapped parameter set, i.e.

$$\mathbf{p}(\mathbf{u}_f') = \mathbf{u}_c' = \text{argmin}_{\mathbf{u}_c \in \Omega_c} r(\mathbf{u}_c, \mathbf{u}_f')$$

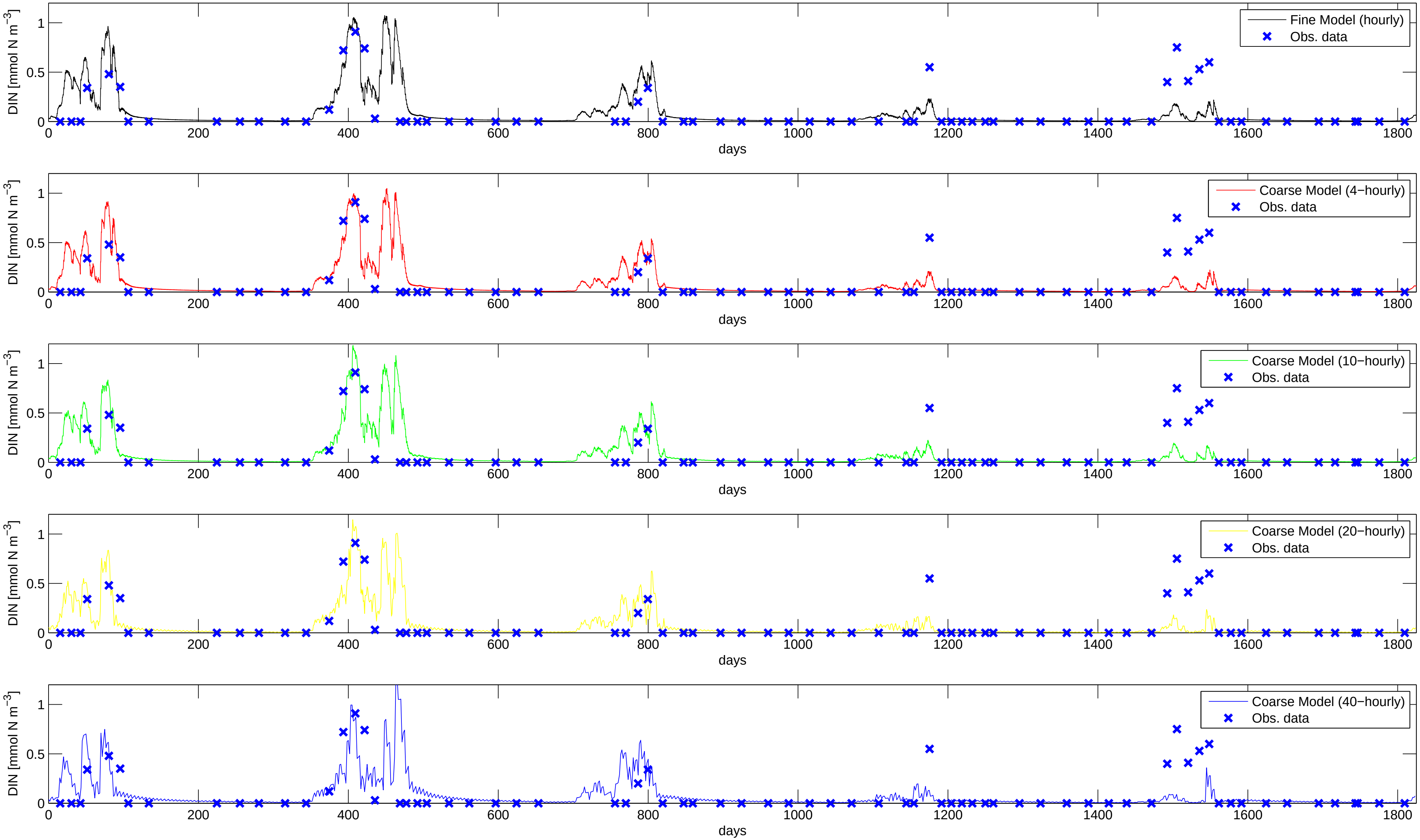
- First approach: using simple steepest descent method

| factor n | $r(\mathbf{u}_f', \mathbf{u}_f')$ | $r(\mathbf{u}_c', \mathbf{u}_f')$ |
|---|---|---|
| 1 | 0 | / |
| 5 | 0.885 | 0.568 |
| 8 | 1.911 | 1.363 |
| 15 | 4.121 | 2.549 |
| 20 | 13.821 | 11.573 |
| 40 | 30.408 | 15.023 |

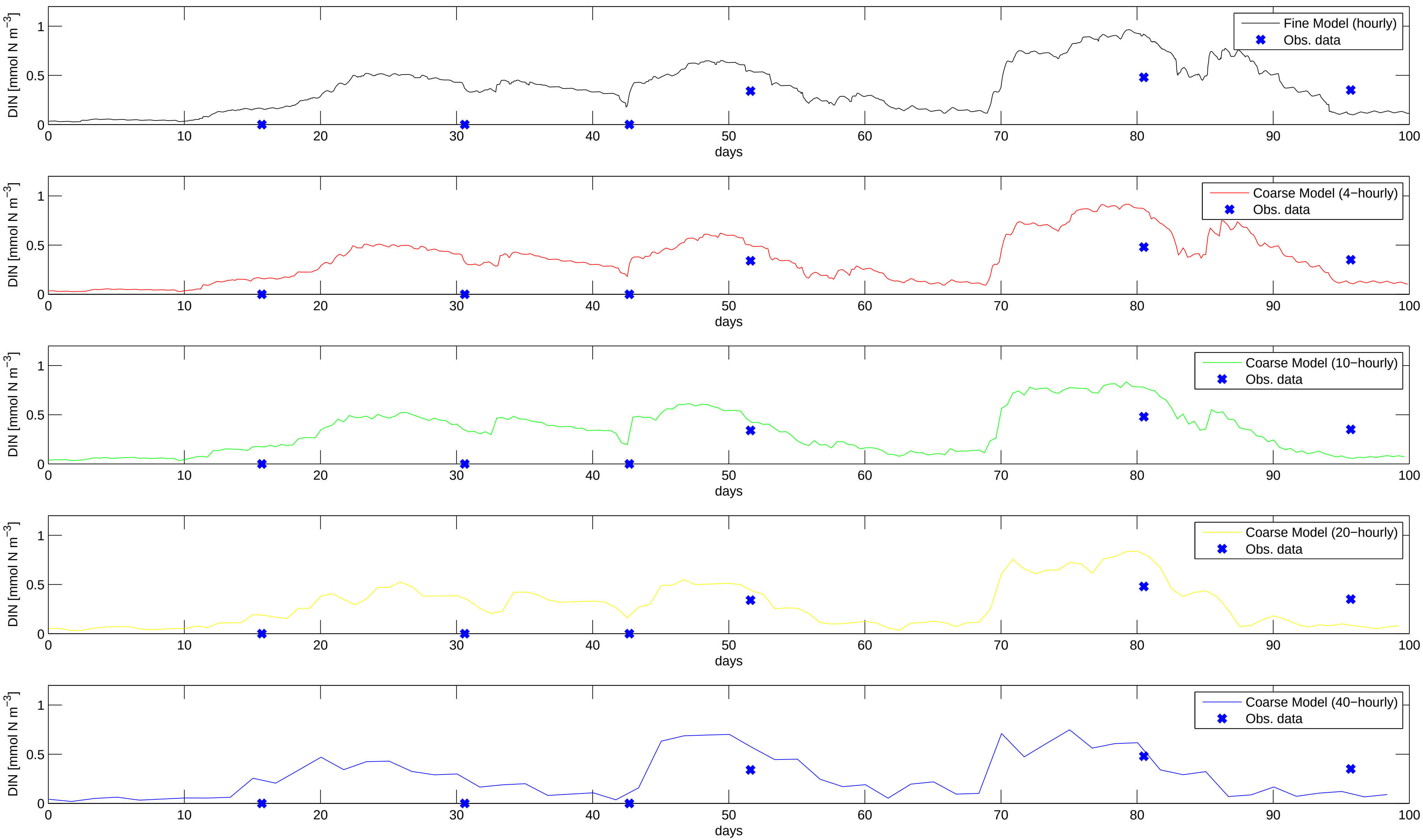
↷ Method seems to be unsuitable to obtain $\mathbf{p}(\mathbf{u}_f)$

↷ Switched over to MATLAB min. toolbox fmincon

# Coarse-Discretization Model



# Coarse-Discretization Model



# Coarse-Discretization Model

$\mathbf{u}_f^* = \text{argmin}_{\mathbf{u}_f} H_f(\mathbf{u}_f) \quad \Leftarrow$ MATLAB fmincon + AD for $J_f$

$\mathbf{u}_c^* = \text{argmin}_{\mathbf{u}_c} H_c(\mathbf{u}_c) \quad \Leftarrow$ MATLAB fmincon + AD for $J_c$

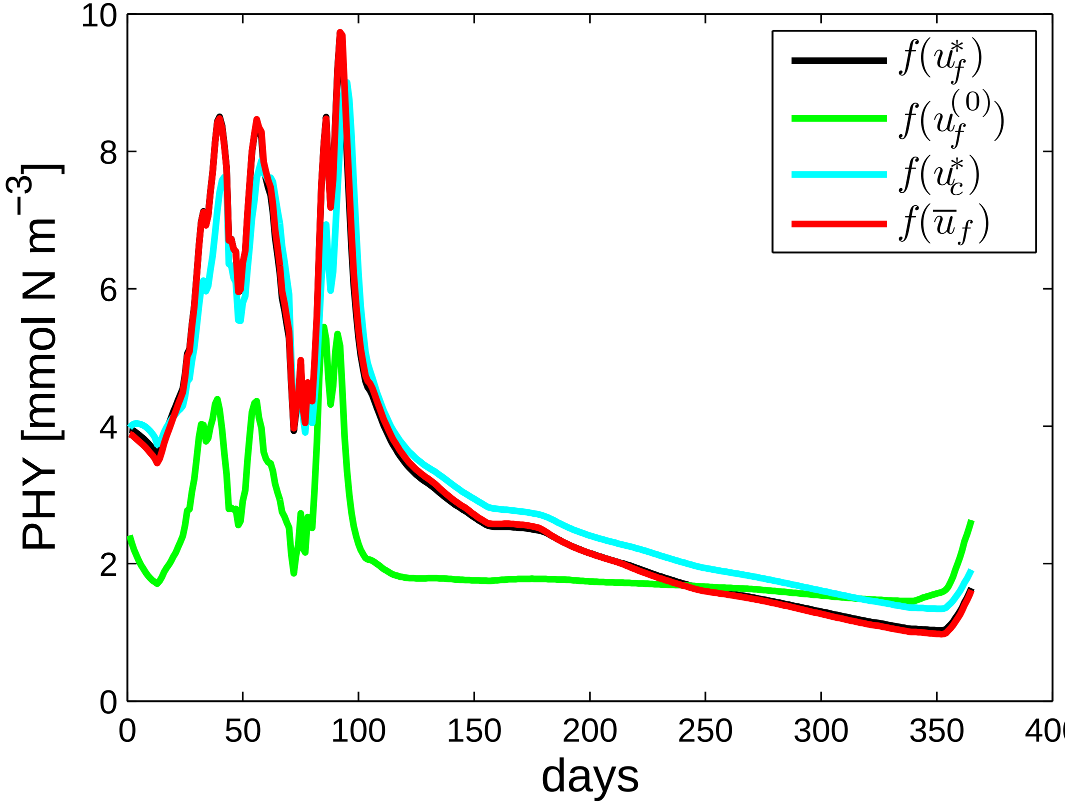
$\mathbf{u}_c = \mathbf{p}(\mathbf{u}_f) = \text{argmin}_{\hat{\mathbf{u}}_c} r(\hat{\mathbf{u}}_c, \mathbf{u}_f) \quad \Leftarrow$ MATLAB fmincon + AD for $J_c$

Primal SM: $\mathbf{F}(\mathbf{u}_f) = \mathbf{p}(\mathbf{u}_f) - \mathbf{u}_c^* \overset{!}{=} 0 \quad \Leftarrow$ Global Quasi-Newton SJN Method

Jacobian of $\mathbf{p}$: $B^{(k)} \approx J_p(\mathbf{u}_f^{(k)}) \quad \Leftarrow$ Broyden rank-one approximation

# Coarse-Discretization Model

| iterate $\mathbf{u}^{(k)}$ | | min $H_f$ | min $H_c$ |
|---|---|---|---|
| $\mathbf{u}_f^{(0)} = \mathbf{u}_c^{(0)}$ | $H_f^{(0)}$ | $1.58 \cdot 10^{-1}$ | $1.58 \cdot 10^{-1}$ |
| | $\delta\mathbf{u}^{(k)}$ | $2.14$ | $2.14$ |
| $\mathbf{u}_c^*$ | $H_f / H_f^{(0)}$ | - | $8.91 \cdot 10^{-2}$ |
| | $\delta\mathbf{u}^{(k)}$ | - | $9.22 \cdot 10^{-1}$ |
| | | | ASM |
| $\bar{\mathbf{u}}_f$ | $H_f / H_f^{(0)}$ | $1.03 \cdot 10^{-3}$ | $3.63 \cdot 10^{-3}$ |
| | $\delta\mathbf{u}^{(k)}$ | $3.65 \cdot 10^{-1}$ | $6.98 \cdot 10^{-1}$ |
| | time | $3683\,s$ | $2722\,s$ |
| | # iter. | $24$ | $37 + 4$ |



# Fourier-Type Model

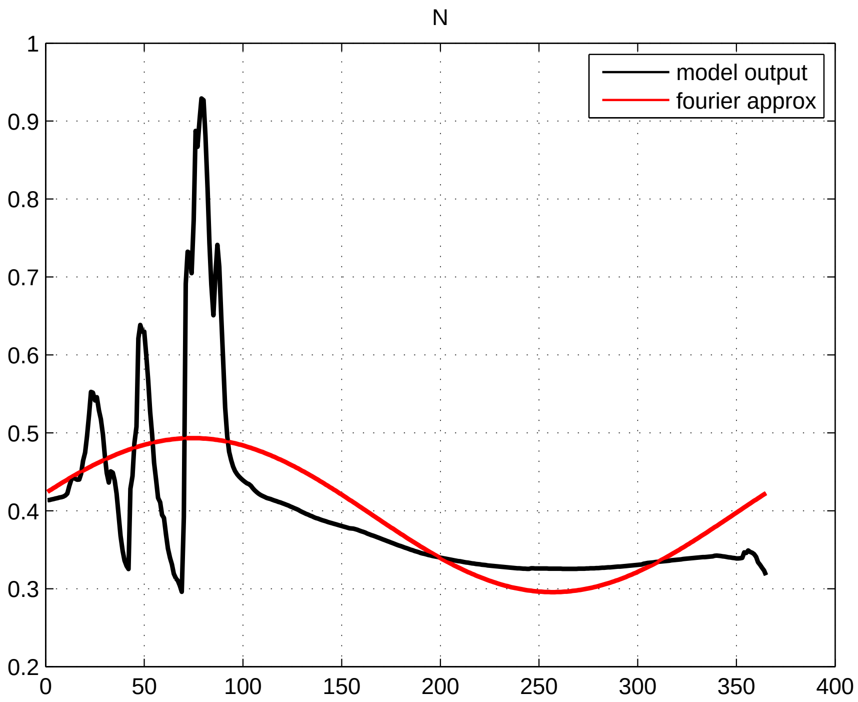


Consider the coarse model $\mathbf{c}$ as the truncated Fourier series at parameters $\mathbf{u}_c$ (= first fourier coefficients)

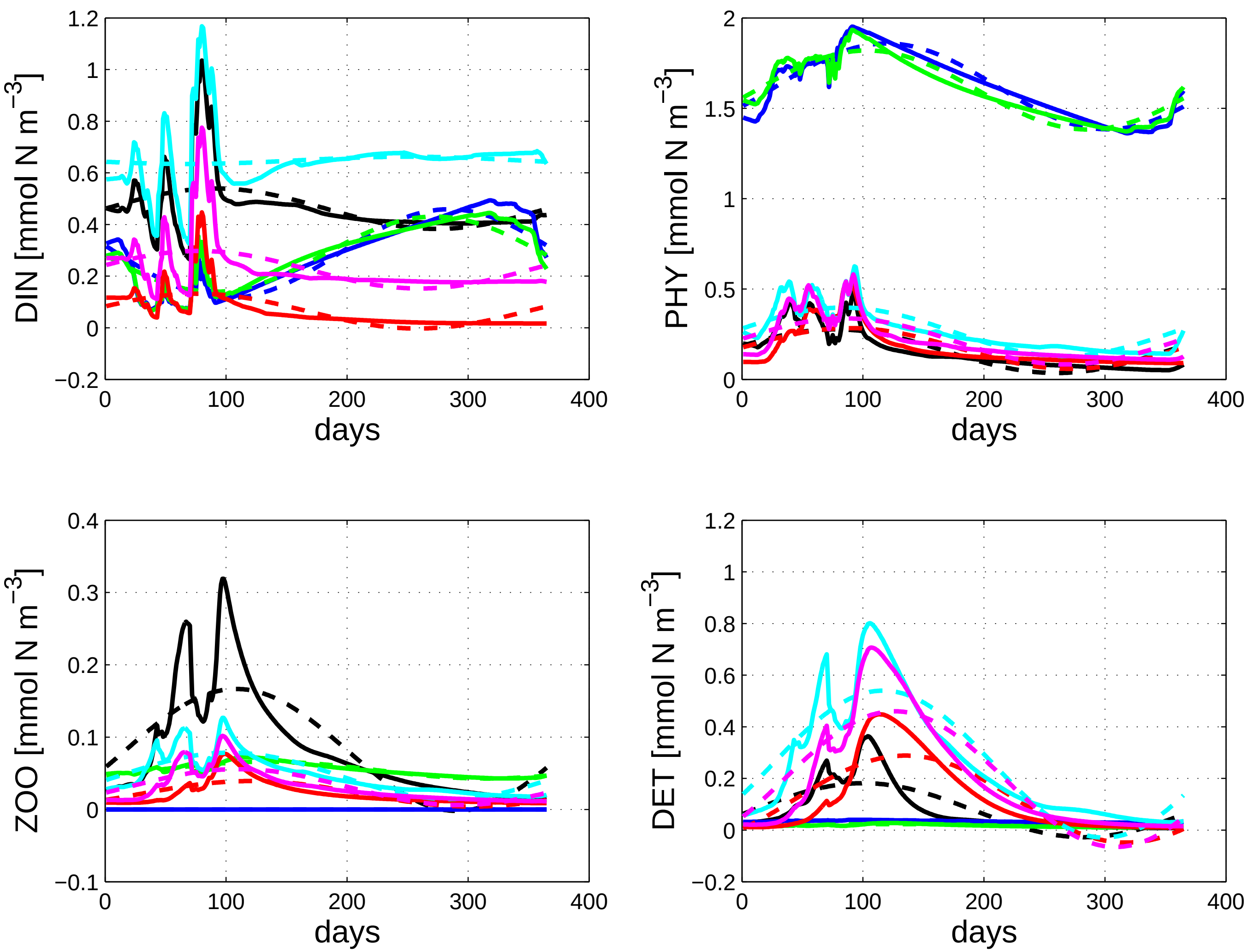
$$\mathbf{u}_c^* = \text{argmin}_{\mathbf{u}_c} H_c(\mathbf{u}_c) = \text{FFT}^{(\text{tr})}(\mathbf{y}_d)$$

$$\mathbf{u}_c = \mathbf{p}(\mathbf{u}_f) = \text{argmin}_{\hat{\mathbf{u}}_c} r(\hat{\mathbf{u}}_c, \mathbf{u}_f) = \text{FFT}^{(\text{tr})}\left[\mathbf{f}(\mathbf{u}_f)\right]$$

$$\mathbf{p} : \mathbb{R}^{12} \mapsto \mathbb{C}^{12} \quad , \quad \mathbf{u}_f \in \mathbb{R}^{12} \quad , \quad \mathbf{u}_c \in \mathbb{C}^{12}$$



# Fourier-type Model




OPT run using ASM. Solid/ dashed lines: fine and coarse model output (one year) $\mathbf{f}(\mathbf{u}_f^{(k)})$, $\mathbf{c}(\mathbf{u}_c^{(k)})$ in iteration $k$ (in the order blue, green, red, cyan, magenta), black: optimal solution $f(\mathbf{u}_f^*), c(\mathbf{u}_c^*)$. Here $\|F^{(5)}\| / \|F^{(0)}\| \simeq 0.06$, $H_f^{(5)} / H_f^{(0)} \simeq 0.04$ .

# Globalized (Quasi-) Newton Method

# Problem Formulation + Algorithms

Suggested (Quasi-) Newton method converges only locally

**Globalization strategy**

Level function : $T(\mathbf{u}_f\,|\,A) := \frac{1}{2}\,\left\|A\,F(\mathbf{u}_f)\right\|^2$

Descent direction : Newton direction $J_F(\mathbf{u}_f)\,\mathbf{s}^{(N)} = -\mathbf{F}(\mathbf{u}_f)$

( since $< \nabla T, \mathbf{s}^{(N)} >_{A=I} = -\mathbf{F}^T J_F\, J_F^{-1}\,\mathbf{F} = -\|\mathbf{F}\|^2 < 0$ )

Linesearch : Find parameter $\sigma$ s.t. $T(\mathbf{u}_f + \sigma\cdot\mathbf{s})\,|\,A) \leq t_k(A)\cdot T(\mathbf{u}_f\,|\,A)$

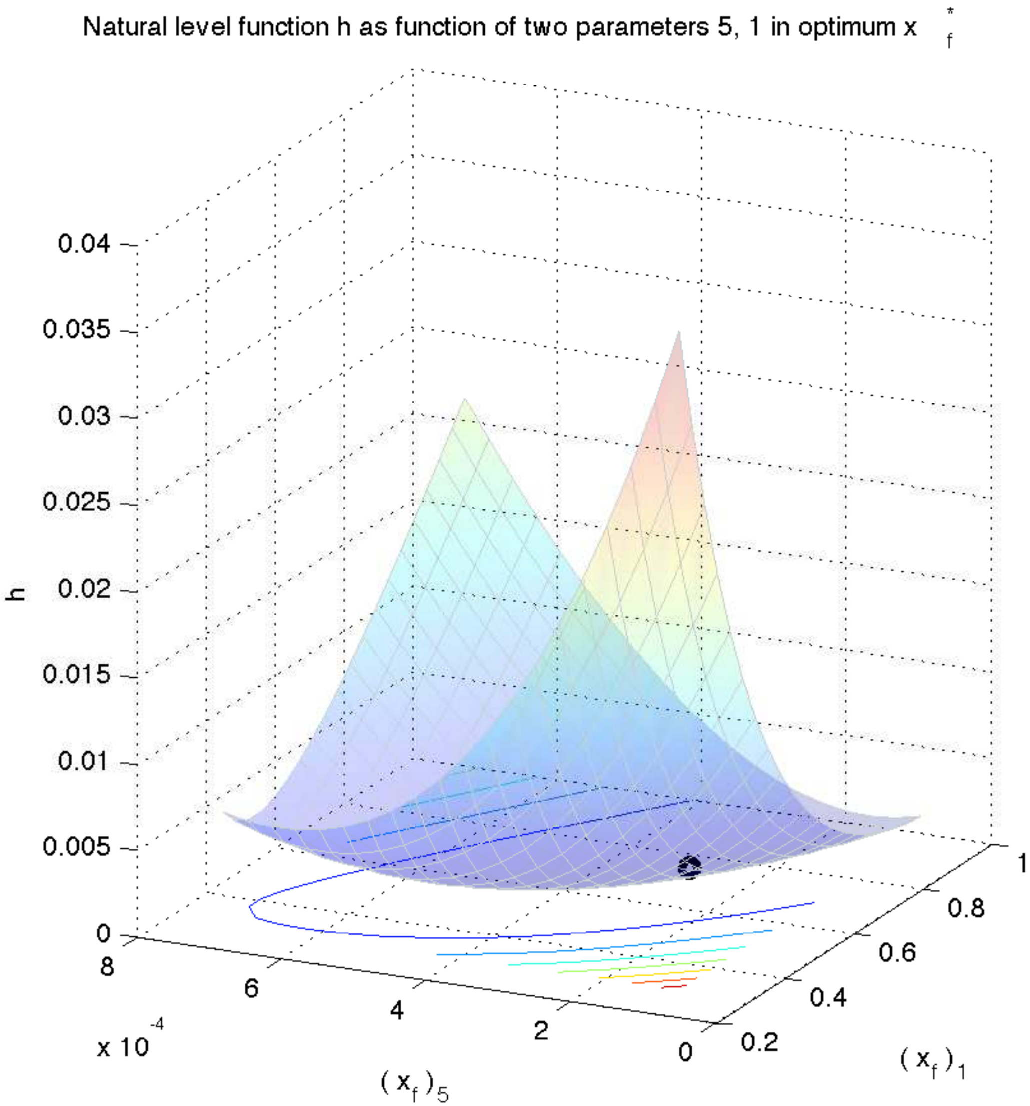
(i) Local minima of the level function $T(\mathbf{u}_f\,|\,A = I)$ where $\nabla T(\mathbf{u}_f) = J_F^T\,\mathbf{F}$, $\mathbf{F} \neq 0$ and $J_F$ singular

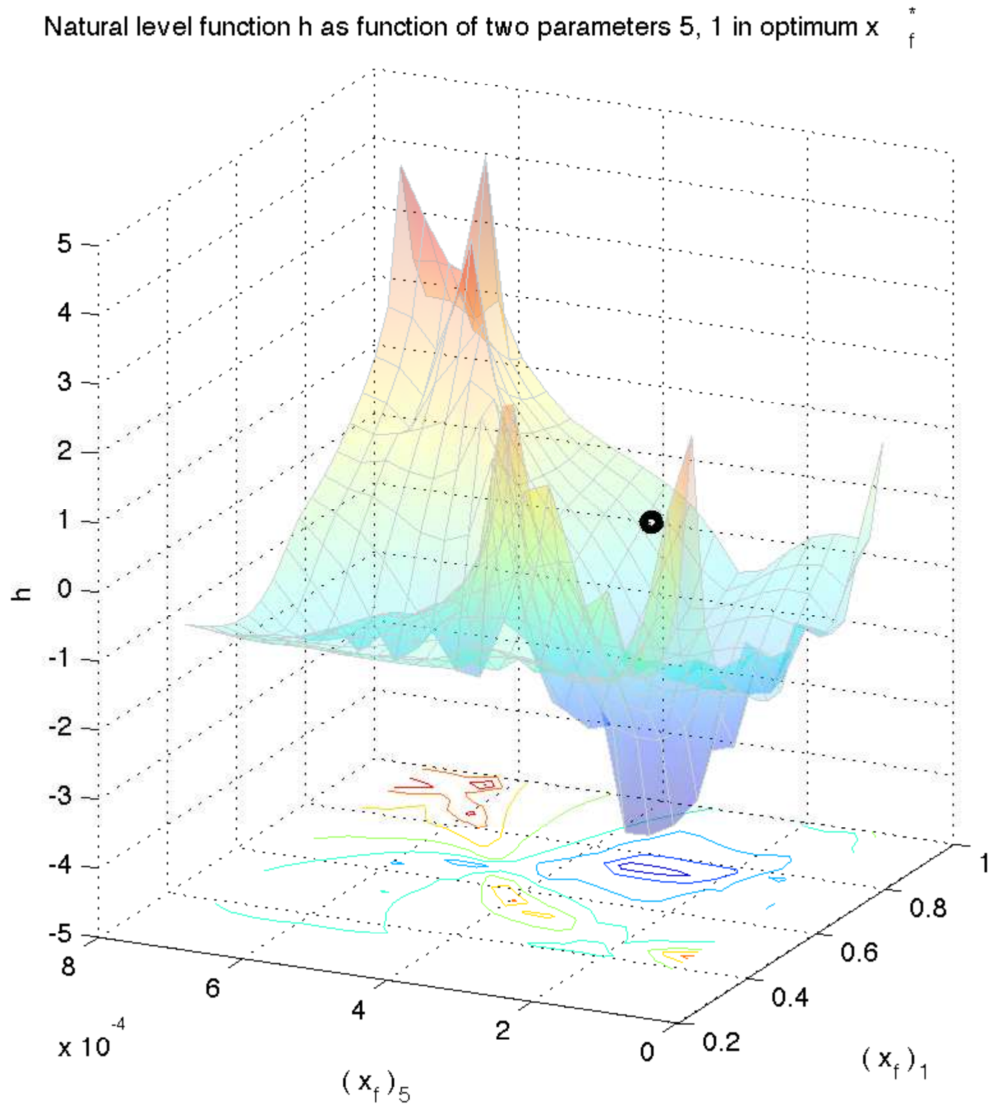
$\curvearrowright$ One can show that the choice $A = J_F^{-1}$ leading to the natural level function is more convenient

(ii) Case $J_f$ ill-conditioned : thus perturbed step $\mathbf{s}^{(N)}$ or $\mathbf{s}^{(QN)}$ might lead to non-descent dir. and breakdown of the algorithm

$\curvearrowright$ Apply rank-strategy (yielding a descent direction for $T(\mathbf{u}_f\,|\,A = J_F^{-1}))$

# Original vs. Natural Level Function



# Simple Testcases



Our focus mainly lies on two global methods:

(i) Global Quasi-Newton SJN method (cf. Kosmol, 1993 )

(ii) Global Newton (locally: Quasi-Newton) method (cf. Deuflhard, 2004 )

Tests with a simple North Atlantic Boxmodel show:

**Method (i)**

$\curvearrowright$ Local min. for $\simeq 40$ % of randomly choosen initial parameters (similar “bad” results for other simple test functions)

**Method (ii)**

$\curvearrowright$ Results follow

**Method (ii) + rank-strategy**

$\curvearrowright$ Results follow

# Open Problems

- Main focus should lie on the development of “appropiate” coarse models
  - Coarse-discretization model $\curvearrowright$ Multigrid methods
  - Linearization of model equations. Where does the focus lie?
  - Furthermore what possible approaches could be done ?
- Suitable validation techniques of the coarse models
- Error analysis
- Where exactly (w.r.t. the algorithms) is reduction of computer time consumption founded ?

- Appropiate SM approaches and involved numerical/ optmization methods
  - Direct (primal through NLE) vs. indirect SM approach (dual through replacing $\mathbf{f}$ by its surrogate $\mathbf{c}[\,\mathbf{p}(\mathbf{u}_f)\,]$) ?
  - Adjoint approach for optimal control of a coarse model?
  - Multipoint PE, implicit SM, other Jacobian approximation, regularization within $\mathbf{p}$,
  - TRASM, Hybrid SM methods

...

Thank you for your attention

[1] John W. B, Qingsha S. Cheng, Sameh A. Dakroury, Ahmed S. Mohamed, Student Member, Student Member, Student Member, Mohamed H. Bakr, Kaj Madsen, and Jacob Søndergaard. Space mapping: The state of the art. 2004.

[2] Mohamed H. Bakr, John, and W. Bandler. Review of the space mapping approach to engineering optimization and modeling. *Optimization and Engineering*, 1:241–276, 2000.

[3] P. Deuflhard. *Newton methods for nonlinear problems: affine invariance and adaptive algorithms*. Springer Series in Computational Mathematics, 2004.

[4] D. Echeverría and P.W. Hemker. Space mapping and defect correction. *Computational Methods in Applied Mathematics*, 5:107–136, 2005.

[5] P. Kosmol. *Methoden Zur Numerischen Behandlung Nichtlinearer Gleichungen Und Optimierungsaufgaben*. Teubner, 1993.

[6] M.F. Pantoja, P. Meincke, and A.R. Bretones. A hybrid genetic-algorithm space-mapping tool for the optimization of antennas. *Antennas and Propagation, IEEE Transactions on*, 55(3):777–781, March 2007.

[7] Markus Schartau. *Data-assimilation studies of marine, nitrogen based, ecosystem models in the North Atlantic Ocean*. PhD thesis, 2001.